

# The Stellenbosch STD Submission

J. A. du Preez    A. S. Visagie    E. S. de Villiers  
A. Strassheim

Department of Electrical and Electronic Engineering  
University of Stellenbosch, South Africa

NIST Spoken Term Detection Workshop, 2006

# Outline

1 System Description

2 Results

3 Discussion

# Databases and Feature Vectors

- Only English CTS
- Call-Home pronunciation lexicon (Pronlex), augmented by in-house letter-to-sound rules
- Trained from total 50 hours from Call-Home, Fisher and Switchboard-1
- Used NIST Dev set for calibration
- Energy normalization and pre-emphasis
- MFCC over 200 Hz–3500 Hz band
- Cepstral mean subtraction
- Group spans of 9 vectors and LDA

# Indexing

- No grammar, context modelling, channel or speaker adaptation
- Train 40 HMM-based phoneme models (3 and 4 states)
- Extract the GMMs internal to these HMMs (256 mixtures)
- Indexing involves scoring speech on these GMMs
- Index files contain compressed phoneme probabilities

# Term Detection

- Each phoneme is a 2 state HMM sharing common conceptual PDF
- The phoneme PDFs directly use the phoneme scores from the indexing file
- Build model for target term by concatenating simple phoneme HMMs
- Background language HMM reflects occurrence probabilities of phonemes in language
- Target locations found via Viterbi algorithm on combined background-target HMM
- Target hypothesis score calculated by normalizing target model score with the background model score

# Our Position

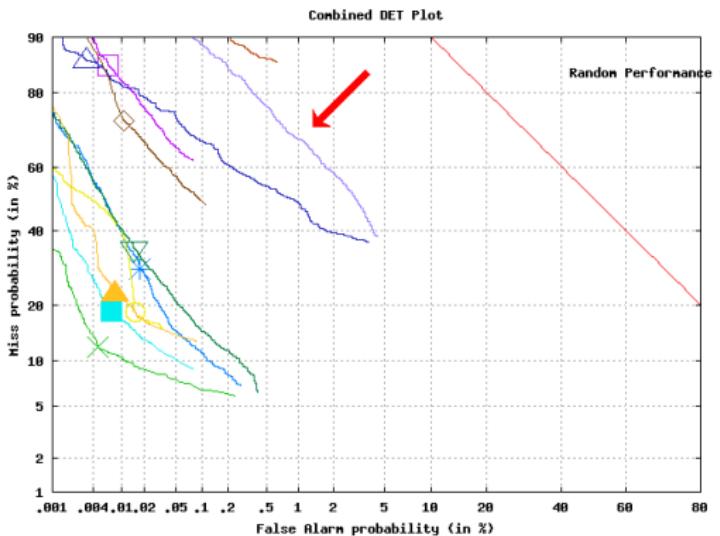


Figure: CTS DET curves

# Score Histograms for Various Target Lengths

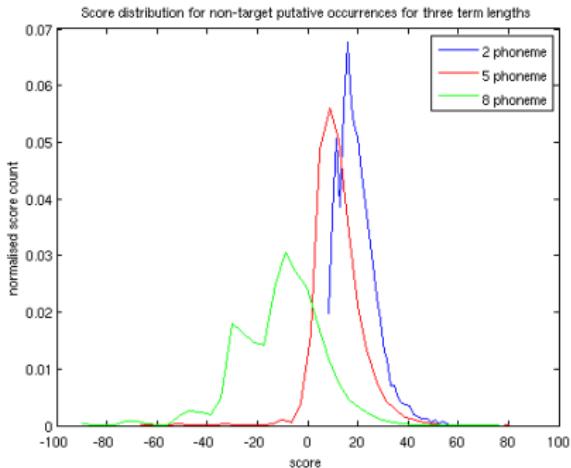
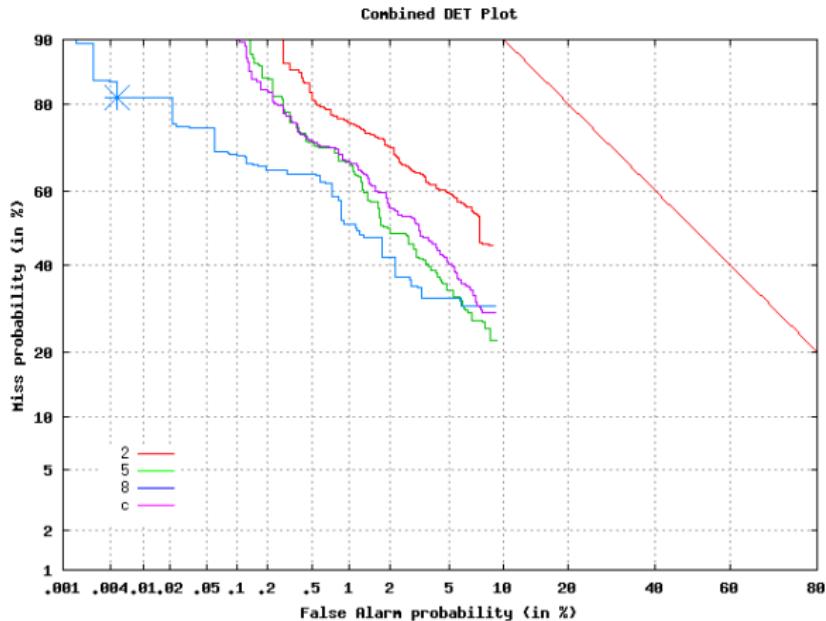


Figure: Histograms of FA scores

# Effect of Target Length on Detection Scores



**Figure:** DET curves broken down by target length

# Conclusions

- Clearly we have a target length calibration problem
- What else?
  - Better post-elimination of FA based on duration?
  - Modelling of phoneme context (CARTs)?
  - Grammars?
  - Larger training databases?
  - Move to word-lattice instead of phoneme-based?

That's it, and thank you. We'll be back!